

## **REMARKS**

Reconsideration of this application is respectfully requested. No new matter has been added.

### **35 USC 103**

The Examiner has rejected claims 1, 3-7, 14, 16-20, 27, and 29-33 under 35 U.S.C. §103 as being unpatentable over Dean et al., U.S. Patent No. 6,321,220 ("Dean") in view of Wical, U.S. Patent No. 6,460,034 ("Wical").

Claims 1, 3-7, 14, 16-20, 27, and 29-33 are patentable under 35 U.S.C. § 103 in view of the references cited by the Examiner. In contrast to applicants, none of the cited references teach (nor does the Office Action cite any portion which even suggests) the pre-identifying of implicitly defined communities including groups of pages of common interest, from a collection of hyper-linked pages, wherein the communities have not been previously identified. Applicants pre-identifying of implicitly defined communities is not in response to a user supplied search query. Applicants invention does not utilize a search query. Applicants invention is actually a broad data mining query of the web graph. The broad data mining query identifies implicitly defined communities which were previously undiscovered.

Dean merely describes a method and apparatus for preventing topic drift in queries in hyperlinked environments. This is in contrast to the pre-identifying of implicitly defined communities including groups of pages of common interest, from a collection of hyper-linked pages, wherein the communities have not been

previously identified. In the scheme described by Dean, for example in flow chart 200 (see Fig. 2 and col. 5-7) a user inputs a query at a search engine. In response to the input query, Dean retrieves matching URLs, prunes those which are not on topic and ranks the remaining nodes. In short, Dean merely addresses the providing of more relevant search results in response to a search query input by a user at a search engine. Dean addresses the issue of topic drift at ll. 22-45 of col. 3:

When a user wants to find web pages related to a particular topic, the user enters a query representing that topic into a search engine. The search engine finds a result set containing a list of web pages relating that topic. Using an algorithm like Kleinberg's algorithm, this result set is expanded to include other pages that are at a predetermined distance from the pages in the original result set. However, the content of these new pages might not be on the same topic as the original query. If pages that are not on the original query are ranked highly, then this is called "topic drift."

Topic drift may occur when using connectivity information to enlarge the size of an initial result set to include other pages that are reachable within a few links of the initial result set because pages that are one or two links away do not always match the given query. Topic drift also may occur as a result of the existence of many mutually reinforcing pages in the result set, for example if the hub and authority pages point to each other.

Thus, a need exists for a method of preventing topic drift in hyperlinked environments when an initial result set is enlarged to include pages that may better match a given user query.

Thus, Dean does not teach or suggest a scheme for pre-identifying implicitly defined communities including groups of pages of common interest, from a collection of hyper-linked pages, wherein the communities have not been previously identified.

Even adding the teachings of Wical does not render the present invention obvious. Wical merely describes a knowledge base research and retrieval scheme in which a search query is utilized. Wical addresses the utilization of a search query in the SUMMARY OF THE INVENTION at l. 41 of col. 2 to l. 15 of col. 3. Thus, Wical does not teach or suggest a scheme for pre-identifying implicitly defined communities including groups of pages of common interest, from a collection of hyper-linked pages, wherein the communities have not been previously identified. Thus, even if the scheme described in Wical were somehow incorporated into Dean, one would still not arrive at the claimed invention. Wical clearly fails to cure the deficiencies noted with respect to Dean, and, therefore, the claims are patentable over the combination of Dean and Wical.

Again, as stated in the previous response, and stated above, none of the cited references teach (nor does the Office Action cite any portion which even suggests) the pre-identifying of implicitly defined communities including groups of pages of common interest, from a collection of hyper-linked pages, wherein the communities have not been previously identified. Applicants pre-identifying of implicitly defined communities is not in response to a user supplied search query.

**Applicants invention does not utilize a search query** (emphasis added).


Applicants invention is actually a broad data mining query of the web graph. The broad data mining query identifies implicitly defined communities which were previously undiscovered.

Accordingly, the present rejections under 35 U.S.C. §103(a) should be removed.

If there are any additional charges, please charge Deposit Account No. 09-0441 .

Respectfully submitted,  
Prabhakar Raghavan, et al.

Date: 5/10, 2004

  
\_\_\_\_\_  
Mark C. McCabe  
Reg. No. 53,291

IBM Corporation  
Almaden Research Center  
Intellectual Property Law C4TA/J2B  
San Jose, CA 95120-6099  
(408) 927-3380